

EXPRESS MAIL LABEL NO.: EL563155634US

DATE MAILED: September 22, 2000

PATENT

INVENTORS: Michael L. EMENS
Reiner KRAFT

SYSTEM AND METHOD FOR ASSOCIATING SEARCH RESULTS

5

BACKGROUND OF THE INVENTION

1. Field of the Invention

The present invention relates to search engines, and more specifically to a system and method for associating search results that were selected by users in previous search sessions.

2. Description of Related Art

The Internet is a massive, complex, and constantly growing environment in which terabytes of information can now be retrieved from computer systems at various locations. However, traversing and navigating the Internet becomes increasingly difficult and more of a time consuming endeavor as this explosive growth continues. Currently, searching for and finding relevant information on the Internet is a difficult task for both beginners and advanced users. Thus, it is becoming more and more important to improve the quality of the information that is returned in response to a user query.

With conventional search engines, a user looking for specific information on the Internet typically submits a query (i.e., a boolean logic construct of words) to the search engine. In response, the search engine uses the query to search the Internet for the desired information, and then returns a list containing zero or more search results based on what was found. Because the Internet is expanding rapidly, these searches are returning larger and larger sets of results. Therefore, the user is required to spend more and more time

EXPRESS MAIL LABEL NO.: EL563155634US

sifting through the results of a search in order to find the information that is actually of interest.

Typically, the list of search results returned by a conventional search engine is sorted by relevance based on how many times the words in the user's query are found in each result. This type of indexing is done automatically based on a formula, and then all of the results returned by the search engine are presented to the user in an order determined by the relevancy scores. While this may be necessary the first time a query is made, a search engine is used by many users and two or more users often will independently search for similar information or topics. The conventional technique of automatically indexing and then listing all results for a query does not take advantage of the superior ability of humans to evaluate information.

Furthermore, recently proposed search result ordering techniques also fail to adequately improve the quality of the results that are returned to the user. For example, it has been proposed to order the results returned by a search engine according to the overall popularity of the results. In particular, each metadata abstract or URL (i.e., possible search result) is assigned a popularity count based on all user selections that have previously been made in response to all previous queries. Then, when a list of search results (i.e., URLs) is returned in response to a particular query, the search engine lists all of the results for the user in an order that is based on the popularity count for each result. In other words, the search results are merely ordered based on each result's overall popularity with users in general. Like conventional search engines, such search result ordering techniques also fail to take advantage of the superior ability of humans to effectively evaluate search results.

Additionally, both the conventional search engines and the recently proposed search result ordering techniques fail to automatically sift through the search results so as to present the user with only the results that are actually of interest. For example, it is common for a typical multiple keyword query of the Internet to return well in excess of

EXPRESS MAIL LABEL NO.: EL563155634US

5

100,000 search results (i.e., URLs). With conventional search engines, at best the user is presented with a well ordered list that contains all of the results found by the search engine in response to the user's query. Conventional search engines lack the ability to sort through the results so as to return a more manageable list to the user, while at the same time effectively providing the user with the information that is actually being sought.

SUMMARY OF THE INVENTION

10

In view of these drawbacks, it is an object of the present invention to remove the above-mentioned drawbacks and to provide a system and method for associating search results based on the search results selected by users in response to previous searches. The system monitors user search result selections relative to search queries. When a search is performed, the user can be presented with those search results that were previously selected by users in response to the same or a similar query. Thus, a user can leverage the knowledge and evaluation time of previous users that were searching for the same or similar information.

15
Another object of the present invention is to provide a system and method for automatically sifting through search results so as to present the relevant or most relevant results. Thus, a smaller and more focused search result list can be presented to the user.

20

One embodiment of the present invention provides a method for associating search results. According to the method, an original list of search results is provided to a first user in response to a first query, and the search results selected by the first user are recorded and associated with the first query. Additionally, a second query that is the same as or similar to the first query is received from a second user, and an alternate list of search results is provided to the second user. The alternate list lists those search results from the original list that have been associated with the first query due to selection by a user. In a preferred method, the original list of search results is displayed to the second

25

EXPRESS MAIL LABEL NO.: EL563155634US

user, a request is received to view the alternate list of search results, and the alternate list of search results is displayed to the second user.

Another embodiment of the present invention provides a system for providing search results. The system includes a search engine, a query database, and a controller. The search engine provides original lists of search results in response to queries, and the query database stores the search results selected by users in response to each of the queries. The controller provides an alternate list of search results in response to another query that is the same as or similar to one of the queries, with the alternate list of search results listing those search results from the original list that have been recorded in the query database as having been previously selected in response to the one query. In one preferred embodiment, the query database stores each selected search result, and for each stored searched result, an alternate results vector for each of the queries in response to which the stored search result was selected.

Other objects, features, and advantages of the present invention will become apparent from the following detailed description. It should be understood, however, that the detailed description and specific examples, while indicating preferred embodiments of the present invention, are given by way of illustration only and various modifications may naturally be performed without deviating from the present invention.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1(a) is a flow chart for an initial query search in accordance with a searching process of a preferred embodiment of the present invention;

Figure 1(b) is a flow chart for subsequent query searches in accordance with the searching process of the preferred embodiment;

Figure 1(c) is a flow chart for providing alternate search results in accordance with the searching process of the preferred embodiment;

EXPRESS MAIL LABEL NO.: EL563155634US

Figure 2(a) is an exemplary illustration of a portion of the search results returned in response to the initial query of Figure 1(a);

Figure 2(b) is an exemplary illustration of a portion of the search results returned in response to the subsequent query of Figure 1(b);

5 Figure 3 is a block diagram of a system for providing search results in accordance with one embodiment of the present invention; and

Figure 4 is a block diagram of an exemplary information processing system for implementing the searching system and process of the present invention.

10 DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Preferred embodiments of the present invention will be described in detail herein below with reference to the attached drawings.

15 To search the Internet with a conventional search engine, a user submits a query to the search engine. The search engine then performs a search and returns one or more pages that list external resources that potentially contain the information being sought by the user. Typically, each external resource (i.e., search result) is identified by a hyperlink (URL) that can automatically take the user to the resource and a short abstract of the resource. The user must read through the list of abstracts returned by the search engine and will eventually activate the hyperlink for one of the search results to allow closer examination of that resource. Later, the user may return to the list of search results and select one or more additional hyperlinks for further examination. Each time the user 20 selects a hyperlink, a page from the corresponding external resource is downloaded into the user's viewing device (e.g., web browser).

25 If a second user later uses the same search engine to independently search for the same information as the previous user (i.e., submits the same query), the second user is also presented with the same list of search results as the previous user. Thus, the second user must also start from scratch and again sift through all of the abstracts of external

EXPRESS MAIL LABEL NO.: EL563155634US

resources. The second user will not be able to take advantage of the time and energy spent by the previous user in evaluating the search results in response to the same query. In other words, conventional search engines force each user to independently evaluate the same search results because each user has no way of knowing the search results that were selected by previous users in response to the same or a similar query.

A searching process according to an embodiment of the present invention in which search results are associated based on user selections in response to similar searches will now be described. In this example, both users are searching for a particular implementation of the Quicksort algorithm. The first user constructs a complex query (e.g., using the "Advanced Powersearch" search form of IBM's jCentral search engine) that is submitted to the search engine. The search engine searches the Internet and returns a search results page that lists 100 hyperlinks (URLs) to external resources that contain various implementations of the Quicksort algorithm.

The user reads through the abstracts for the resources and eventually chooses the hyperlink for one of the resources that appears to contain the desired information. The user is taken to the selected external resource and a page is downloaded into the user's viewing device (e.g., web browser) to allow the user to more closely examine the information provided by the selected search result. After reviewing this resource, the user may return to the original search results page and select one or more of the other hyperlinks to external resources for further review. As the first user peruses the search results returned by the search engine, all of the hyperlinks that are selected for further examination are recorded. For example, in one embodiment every time the first user chooses one of the hyperlinks on the results page, that eternal resource (URL) is associated with the user's original search query and this relationship is stored in a database.

Later, a second user seeking the Quicksort algorithm submits the same query to the search engine. In response, the search engine retrieves a list of search results that

EXPRESS MAIL LABEL NO.: EL563155634US

contains hyperlinks (URLs) to external resources. Before presenting the search results to the second user, each hyperlink is cross referenced with the selections made by other users that submitted the same query to the search engine. More specifically, the database is searched to identify all of the retrieved hyperlinks that have previously been chosen for the same query. For each search result that was previously chosen, the search engine places an icon (e.g., labeled "See What Others Selected" or "Show Alternate Results") next to that search result when displaying the list of search results to the second user.

At this point, the second user can peruse the search results in the normal manner or can select one of the alternate results icons. If an alternate results icon is selected, the search engine returns a new list to the second user that lists those search results (URLs) that have previously been selected in response to the same query (as indicated by the relationships stored in the database). Thus, the second user can take advantage of the time and energy that previous users seeking the same information (such as the first user) used to evaluate the search results. In general, this will present the second user with a much smaller and more manageable list of more relevant search results.

While one specific embodiment of the invention has been described above, many modifications are possible for specific applications. For example, a count of the number of times each hyperlink has previously been selected in response to the same query can also be stored in the database. The relative counts can then be used to order the original list of search results and/or the alternate list of search results that were previously selected. Further, the actual count values can be displayed to the user along with the search results in order to further assist the user in making a selection. The count value can also be used to set a minimum threshold of selections before a search result is added to the alternate list. Similarly, the alternate results icon or link can be displayed in various ways and at various times. For example, the icon can be presented alongside some or all of the search results, at the beginning and/or end of the original list of search results, or on each page of the original list of search results. Additionally, the alternate

EXPRESS MAIL LABEL NO.: EL563155634US

results icon can be displayed before, at the same time as, or after the search engine displays the original list of search results.

Furthermore, the search engine may contain various other features to assist the user in finding the desired information. For example, the search engine may be able to identify similar but different queries that would return the same (or virtually identical) search results using conventional query matching techniques. This would allow the search engine to cross reference the search results with the selections made by previous users that submitted similar but different queries to the search engine. Thus, a user could benefit from the selections made by previous users that were seeking the same or similar information using different queries. Additionally, the search engine could contain query refinement features to assist users in formulating queries. Then, a user can submit a partial query string that the search engine will enhance based on which search results the user selects or the abstracts of the search results that are selected. For example, if the user selects a hyperlinked word in one of the abstracts, then that word is concatenated to the previous search query to form a more complete query string.

A searching process according to a preferred embodiment of the present invention will now be explained in more detail with reference to the flow charts of Figures 1(a), 1(b), and 1(c). Figure 3 shows a system that performs the searching process. As shown, the system includes a controller 10, a search engine 20, a repository 22, and a query database 24. The components of the controller 10 include a session manager 12, a query manager 14, an alternate results manager 16, and an associator 18. The session manager 12 intercepts all user input so as to allow the controller 10 to perform the searching process of the preferred embodiment.

Initially, there are no entries in the query database 24. When a first user submits a first query using an input device 26, the session manager 12 determines if a session ID has been established (step S10). If not, the session manager 12 creates a session ID (step S12) for the user using standard web technologies (e.g., cookies or active server pages) to

EXPRESS MAIL LABEL NO.: EL563155634US

store various state and session information. The session ID allows subsequent queries to be matched with the search results selected by this user. After a session ID has been established, the session manager 12 submits the first user's query to the search engine 20. The search engine 20 performs a search and returns the search results to the session manager 12 (step S14). Then, the controller 10 operates to display a list of the search results to the user through an output device 28 (step S16).

For the initial query, it is immaterial whether the search results are formatted in a standard manner or by using the process described below because the query database is empty. Figure 2(a) shows an exemplary list of search results presented to the first user in response to the query "cats and dogs". The first user then peruses the list of search results and may select one or more for further review (step S18). In the controller 10, each selection made by the first user is intercepted by the session manager 12 and routed to the associator 18 in the form of the session ID, the query string, and the selected search result. Because the query database is empty, the associator 18 creates an entry in the query database 24 and creates an "Alternate Results Vector" (ARV) for the query string (step S19).

In this embodiment, the alternate results vector is simply a single element vector composed of a query string and an associated search result (URL) list. Each URL keeps a list of query strings used to get to that URL, and each of those query string entries keeps a list of other URLs also in that pool. This allows the system to handle any number of quantifiers for resource relevance or irrelevance. At this point, all of the search results that were selected by the first user have been effectively recorded so that an alternate list of search results can be displayed on subsequent submissions of the same or a similar query string.

Later, when a second user submits a query string, the session manager 12 again determines if a session ID has been established (step S20). If not, the session manager 12 creates a session ID (step S22) for the user. After a session ID has been established, the

EXPRESS MAIL LABEL NO.: EL563155634US

session manager 12 submits the first user's query to the search engine 20, which performs a search and returns the search results to the session manager 12 (step S24). The session manager 12 then forwards the search results along with the session ID and query string information to the alternate results manager 16. The alternate results manager 16 pairs each search result with the query string and submits this information to the query manager 14, which searches for the search result in the query database 24. If no entry exists for that search result, a false status is returned to indicate that there are no alternate results.

On the other hand, if an entry exists for the search result, a search is then made for an Alternate Results Vector (ARV) for the query string. If an ARV for the query string is not found, a false status is returned to indicate that there are no alternate results.

Conversely, if an ARV for the query string is found, the query manager 14 returns a true status to indicate that there are alternate results (step S26). Then, the alternate results manager 16 formats the list of search results for display on the output device 28. In particular, alongside each search result for which a true status was returned, there is displayed a graphical element such as an alternate results button to indicate to the user that an alternate list of search results is available (step S28). For example, Figure 2(b) shows a list of search results displayed to the second user in response to the same query if the first user selected search results (1) and (3) of Figure 2(a).

When the list of search results is displayed, the second user can either select one of the search result for further review or select the alternate results button. If one of the search results is selected, the browser goes directly to that page (or the request is routed through the session manager). Even if the alternate results button is not selected, the mere displaying of the button alongside each of the search results that was previously selected by the first user assists the second user by indicating which of the currently displayed search results had been previously selected.

On the other hand, if the alternate results button is selected, the session manager 12 forwards the session ID, query string, and alternate results message to the alternate

EXPRESS MAIL LABEL NO.: EL563155634US

results manager 16. The alternate results manager 16 then references the results from the query database 24 using the query string. Next, all of the search results in that ARV are retrieved directly from the search engine (step S30). All ARV results found in the search engine are then formatted for display on the output device (steps S32 and S34). Thus, the 5 second user is presented with alternate list of results that matches the query string that was submitted.

Accordingly, the search system and method of the present invention takes advantage of the superior ability of humans to evaluate search results. In particular, the typical user selects a search result after evaluating its abstract and deciding that this 10 search result is likely to contain the information being sought by the query that was submitted. This human knowledge is automatically collected to organize search results for a particular query and can later be reused by other users. Because users can view and select from prior user selections, they save time reading abstracts. This integration of human evaluation capabilities makes browsing drastically improved, simplified, and more user friendly. Furthermore, the search system and method of the present invention makes browsing easier and easier for each subsequent user.

The system and technique of the present invention can be implemented in hardware, software, or a combination of the two. For example, at least a portion of the system can be embodied in software programs that are stored on a computer-readable medium (e.g., magnetic disk, optical disk, or non-volatile memory) for execution by a computer. Similarly, such processes can be embodied in a web server and client like those illustrated in Figure 4. Additionally, embodiments of the present invention described above relate to Internet search engines. However, the system and method of the present invention are not so limited and could be used for other search functions and with 20 any type of network. Likewise, embodiments of the present invention can be written in Java, C, or any other language. Other design choices, such as front end of the query tool, the format of the displayed search results, and the processes for retrieving the search 25

EXPRESS MAIL LABEL NO.: EL563155634US

results and transferring information could also easily be adapted. Further, embodiments of the present invention may not include all of the features described above.

While there has been illustrated and described what are presently considered to be the preferred embodiments of the present invention, it will be understood by those skilled in the art that various other modifications may be made, and equivalents may be substituted, without departing from the true scope of the invention. Additionally, many modifications may be made to adapt a particular situation to the teachings of the present invention without departing from the central inventive concept described herein. Therefore, it is intended that the present invention not be limited to the particular embodiments disclosed, but that the invention include all embodiments falling within the scope of the appended claims.

U.S. PATENT AND TRADEMARK OFFICE